http://dx.doi.org/10.7494/human.2019.18.1.95

Jacek Neckar*

Jagiellonian University
ORCID: 0000-0002-6821-1651

Magdalena Lazaro Szlachta**

Jagiellonian University
ORCID: 0000-0002-0377-6387

GENDER DIFFERENCES IN EMPATHIZING-SYSTEMIZING: THE INFLUENCE OF GENDER STEREOTYPE AND SOCIALLY DESIRABLE RESPONDING

The main aim of the present studies was to investigate the influence of social and cultural factors on gender differences in empathizing-systemizing. Study 1 was designed to control for the socially desirable responding in gender differences in empathizing-systemizing. In Study 2 we wanted to investigate whether the activation of gender stereotype would influence gender differences in the questionnaire and the ability test that measured empathizing-systemizing. Consistently across our two studies and the two measurement methods used (the questionnaire and the ability test), women scored higher in empathizing and the size of the effect was medium. Socially desirable responding had no effect on the size of gender differences in empathizing. However, the activation of the gender stereotype made respondents, especially women, present themselves as more empathetic persons. In addition, the stereotype activation produced a performance boost on the systemizing ability test in men, whereas no effect was observed in women.

Keywords: empathizing-systemizing, gender stereotype, socially desirable responding

INTRODUCTION

Although the historical underpinnings of an empathy construct go back to German psychology of the 19th century, recent decades have witnessed a renewal of interest in empathy (e.g. de Waal 2008; Iacoboni 2009). The substantial number of recent papers have stressed the evolutionary origin of empathy rather than its cultural roots. Research of de Waal (2008, 2009), for instance, reveals empathy as a mechanism common to many species of mammals, not only

* Corresponding author: Jacek Neckar, Uniwersytet Jagielloński, Instytut Psychologii, ul. Ingardena 6, 30-060 Kraków; e-mail: jacek.neckar@uj.edu.pl.

** Corresponding author: Magdalena Lazaro Szlachta, Uniwersytet Jagielloński, Instytut Psychologii, ul. Ingardena 6, 30-060 Kraków; e-mail: szlachtamagda@gmail.com.

primates. The current research focuses on testing Baron-Cohen's empathizing-systemizing theory of gender differences in cognition (Baron-Cohen 2002).

The construct of empathy has been traditionally regarded as multidimensional and defined as (i) cognitive ability to know the mental states of others, (ii) ability to experience emotional reactions that are appropriate to the mental state of others, and (iii) ability to act upon this knowledge and emotions, e.g. consolation of the distressed or helping the suffering (Bateson 2009). Among various attempts to capture the multidimensionality of empathy, Baron-Cohen's construct of empathizing stands out as a well-grounded proposal (Baron-Cohen 2009; Baron-Cohen and Wheelwright 2004). It has been developed as a part of the empathizing-systemizing theory (Baron-Cohen 2009). The theory explains both the autism spectrum disorder and gender differences in empathizing-systemizing. Empathizing encompasses the aforementioned multidimensionality of empathy, but is also related to the theory of mind and taking the intentional stance (Baron-Cohen 2005). Empathizing consists of two components: the ability to attribute mental states to other people and the desire to respond to those mental states with appropriate emotions and actions (Baron-Cohen 2004; Baron-Cohen and Wheelwright 2004). Systemizing is defined by Baron-Cohen as a drive to analyse or create systems, i.e. a rule-governed body of knowledge that can refer to mechanical, abstract or numerical domains (Baron-Cohen 2003, 2006).

Emotional intelligence, which has generated an extensive amount of research, is a construct related to empathizing. The original idea of emotional intelligence (Salovey and Mayer 1990) was that some persons have the ability to recognize and use emotions to enhance their thinking and action better than others. Yet other researchers have conceptualized emotional intelligence in broad and eclectic ways and included not only abilities but also dispositional traits such as self-esteem, optimism or self-management (Bar-On 2004; Petrides and Furnham 2001). It seems that the approach to emotional intelligence that is closest to the construct of empathy is the original one of Salovey and Mayer: the ability approach. Research on emotional intelligence particularly important for this study relates to gender differences in this domain. A recent study performed on a large sample of adults supported the idea of gender differences in ability emotional intelligence, i.e. women tend to outperform men, although the size effect is rather small (Cabello et al. 2016).

There is evidence suggesting that not only empathy but also the empathizing-systemizing dimension is related to biological factors (Baron-Cohen 2006, 2009). According to the empathizing-systemizing theory of Baron-Cohen (2002, 2009), persons with autism spectrum disorder lie at the extreme end of both empathizing and systemizing normal distribution (i.e. very low empathizing, and high systemizing). Since autism spectrum disorder is more prevalent in males compared to females, this pattern of low empathizing and high systemizing has been referred to as an extreme male brain hypothesis of autism (Baron-Cohen 2002). In a recent study (Baron-Cohen et al. 2014) performance of individuals with autism showed no gender difference in a test of the cognitive aspect of empathy, whereas females performed significantly better on this test compared to males in the healthy control group. This result corroborated the extreme male brain hypothesis of autism.

Related to the assumption of biological foundations of empathizing-systemizing is the notion that there are gender differences in that domain. According to Baron-Cohen (2003)

women score higher on empathizing and lower on the systemizing dimension when compared to men. This difference has been confirmed in numerous studies (Baron-Cohen, Knickmeyer and Belmonte 2005; Baron-Cohen and Wheelwright 2004; Goldenfeld, Baron-Cohen and Wheelwright 2005). Baron-Cohen (2003) proposes that the female brain is predominantly hard-wired for empathy or social communication skills, whereas the male brain is hard-wired for focusing on how the systems work and trying to build them (Baron-Cohen 2003). Although the exact biological mechanisms that underlie empathizing-systemizing are still to be discovered, there are some data that point to the biological nature of gender differences on these dimensions (Chakrabarti, Bullmore and Baron-Cohen 2006; Knickmeyer, Baron-Cohen, Raggatt and Taylor 2006). An important way of showing biological influence on gender differences is to demonstrate that these differences are constant across different cultures. In an important cross-cultural study, in which four different samples from Malaysia, Slovenia, Switzerland and Turkey were compared (Zeyer et al. 2013), a very stable effect of systemizing on motivation to study science was detected across all four cultures. Their structural model showed that gender only indirectly, via systemizing, influenced the motivation to study science, but the effect size of gender differences in empathizing-systemizing was similar to that obtained in European samples (Groen et al. 2015).

Only a few studies conducted so far have shown that gender differences in empathizing--systemizing are not entirely immune to cultural influences (Berthoz, Wessa, Kedia, Wicker and Grezes 2008; Groen, Fuermaier, Heijer, Tucha and Althaus 2015; Preti et al. 2011; Zeyer et al. 2013; Zheng and Zheng 2015), especially when participants from the West and Far East are compared. These studies suggest that gender differences in empathizing are comparable in European samples (Groen et al. 2015; Preti et al. 2011) and similar in a Japanese sample (Wakabayashi et al. 2007), whereas there were no gender differences in empathizing in a Korean sample (Kim and Lee 2010). Similarly, recent data from a Chinese study (Zheng and Zheng 2015) revealed no gender difference in empathizing, which suggests an existence of culturally determined patterns. In addition, systemizing has shown more consistent and stronger gender differences in Asian populations (Wakabayashi et al. 2007; Zheng and Zheng 2015) compared to European ones (Baron-Cohen et al. 2003).

There are several aspects of social and cultural factors that could potentially influence empathizing-systemizing (Baron-Cohen 2006; Baron-Cohen et al. 1996; Wakabayashi et al. 2007). One of them that has not been investigated so far in relation to empathizing-systemizing is the stereotype threat. An important form of stereotype is gender stereotype, which refers to a set of generalized characteristics that apply only to men or women. The threat of gender stereotype has been documented to decrease or enhance the performance of women and men on various measures (Aronson and Steele 2005; Steele 1997; Wheeler and Petty 2001), including behavioural and cognitive tasks (Grand, Ryan, Schmitt and Hmurovic 2011; Seibt and Forster 2004) and questionnaires (von Hippel et al. 2005). Since empathizing is closer to a stereotypical view of female skills and systemizing is linked more to male technical talents, one could expect an influence of gender stereotype on self-perceived abilities in empathizing-systemizing.

Over the past twenty years research has suggested that the mere existence of negative stereotypes is enough to create an intellectual environment which undermines the performance

of those stigmatized. This negative effect has been named 'stereotype threat' (Steele 1997) and is defined as a situational predicament in which persons are at risk, through their actions, of confirming negative stereotypes about their groups (Steele and Aronson 1995). The research on stereotype threat suggests that when a person who is aware of the negative stereotype impugning his or her intellectual ability is in a situation that requires him or her to show that ability may fear confirming the stereotype, and that fear may jeopardize his or her performance (Schmader, Johns and Forbes 2008). More specifically the possible mechanisms that compromise the performance of a person under stereotype threat are decreasing working memory, ruminating on one's own ability, and unsuccessful efforts to regulate one's thoughts and emotions (Schmader, Johns and Forbes 2008). A parallel to stereotype threat is the idea of 'stereotype boost', which has been conceptualized as an influence of positive stereotypes on a person's actions that can result in increasing his or her performance (Shih, Pittinsky and Ho 2011). In the case of stereotype boost, possible mechanisms that facilitate the performance boost are: decreasing of anxiety related to the task, increased expectations, increased efficiency in neural processing, and greater persistence in task completion (Shih, Pittinsky and Ho 2011).

Another social factor that has not been controlled in the context of empathizing-systemizing but introduced mainly at the reliability and validity check phase of the empathizing-systemizing questionnaires, is socially desirable responding. Self-reports of empathizing-systemizing can potentially be contaminated by social desirability and self-favouring biases. As both empathizing and systemizing scales contain items that have clearly positive, i.e. socially desirable, interpretation and, what is more, some of these items are more positive for women and some are more positive for men, gender differences in empathizing-systemizing can be contaminated. The four studies in which correlation analysis was performed between the desirable responding scores and different measures of empathy, including empathizing (Baldner and McGinley 2014; Lawrence, Shaw, Baker, Baron-Cohen and David 2004; Preti et al. 2011; Trent, Park, Bercovitz and Chapman 2016) have shown a significant (although weak) relationship between them, which suggests that controlling for socially desirable responding may influence the empathizing-systemizing scores in general, and gender differences in these dimensions in particular.

In addition, in all the aforementioned studies the socially desirable responding has been exclusively measured with the Social Desirability Scale (Crowne and Marlowe 1960). However, there has been an interesting development in the area of socially desirable responding since Marlowe and Crown's original work (see Paulhus 1984; Paulhus and Reid 1991; Roth, Harris and Snyder 1988; Sackeim and Gur 1978). For example, research has revealed the existence of two separate aspects of socially desirable responding: a conscious tendency to present oneself in a favourable light, and a more implicit tendency to do so (Paulhus 1984, 1998). Factor analytic studies have consistently supported the existence of the two factors, labelled Alpha (Block 1965) and Gamma (Wiggins 1964). Paulhus (1984, 1986) provided the most extensive evidence of the existence of these two factors, which he labelled 'self-deception' and 'impression management', respectively. The most widely validated current measure of socially desirable responding which captures the two factors is the Balanced Inventory of Desirable Responding (Paulhus 1989). It has become a common and

recommended measure of the two-factor model of desirable responding (see meta-analysis by Li and Bagger 2006).

The main aim of the present studies was to investigate the possible moderating influence of social and cultural factors on gender differences in empathizing-systemizing in healthy participants. The rationale behind this aim was that testing hypotheses of possible biological/evolutionary origins of psychological mechanism seldom focused on alternative explanations, i.e. cultural factors (see Simpson and Campbell 2016, for a similar argument). We take a Popperian stance in this study, i.e. the best way to show the truth of a hypothesis is to fail to falsify it. If the gender differences in empathizing-systemizing are not explained by social and cultural factors, the hypothetical biological or evolutionary roots of the differences can be treated as probable. Study 1 was designed to control for the socially desirable responding in gender differences in empathizing-systemizing, in order to exclude this factor as a possible explanation of gender differences. To our best knowledge, socially desirable responding has not been controlled in any published study that examined gender differences in empathizing-systemizing. Also, the Balanced Inventory of Desirable Responding (Paulhus 1989) has not been used so far to analyse the relationship between the empathizing-systemizing questionnaires and socially desirable responding. In Study 2 we wanted to investigate whether the activation of gender stereotype would influence gender differences in performance in the domain of empathizing and systemizing measured by self-report and ability instruments. Although a strong effect of gender stereotype has been documented in many domains (see Aronson and Steele 2005 for a review), there has been no attempt so far to test its influence in the empathizing-systemizing domain. If confirmed, the gender stereotype effect on performance in empathizing-systemizing can have important implications on the diagnostic validity of the instruments that measure empathizing-systemizing for detecting persons on the autism spectrum. Also, the gender stereotype effect can suggest greater caution in corroborating sex differences that have allegedly biological or evolutionary origins.

STUDY 1

To control for socially desirable responding in gender differences in empathizing and systemizing, we administered questionnaires developed by Baron-Cohen and collaborators (Baron-Cohen and Wheelwright 2004; Wheelwright et al. 2006) and Paulhus' Balanced Inventory of Desirable Responding (Paulhus 1989). We expected that desirable responding would influence gender differences, and therefore controlling for desirable responding would reduce the effect size of gender differences. We expected this reduction to be rather small. Our prediction of a small decrease in gender differences was based on the presumed biological foundations of gender differences in empathizing and systemizing (Baron-Cohen 2009). It was also based on significant but low correlations between empathizing-systemizing and socially desirable responding revealed in the studies in which Baron-Cohen and collaborators' questionnaires of empathizing/systemizing and the Marlowe-Crowne Social Desirability Scale were used (Baldner and McGinley 2014; Lawrence, Shaw, Baker, Baron-Cohen and David 2004; Preti et al. 2011; Trent, Park, Bercovitz and Chapman 2016).

METHOD

PARTICIPANTS

A total of 93 volunteers (60 women) were recruited from various departments of the Humanities of the Jagiellonian University. They were given extra credit for their participation. The mean age of the participants was 21.3 years, with standard deviation 1.89, and it ranges between 19 and 29.

MATERIALS AND PROCEDURE

Empathy Quotient (EQ, Baron-Cohen and Wheelwright 2004) was used to assess empathizing. The original Empathy Quotient was a self-report questionnaire with 60 items, including 20 filler items. However, the version that is currently recommended on the Cambridge website uses only 40 interpretable items. Empathy Quotient has high reliability and validity indices (Baron-Cohen and Wheelwright 2004; Lawrence, Shaw, Baker, Baron-Cohen and David 2004), with α Cronbach equalling .92 and test-retest correlation equalling .84. For the purpose of the present research a Polish version of the 40-item EQ (without filler items) was prepared. The translation was checked by a bilingual specialist in psychology. A 4-point Likert-type scale was used in the questionnaire. Participants received 1 or 2 points for an "empathic" response (slightly agree and strongly agree, respectively) and 0 for the two other responses. The Cronbach α for the Polish version of Empathy Quotient was .81.

Systemizing Quotient (SQ, Baron-Cohen, Richler, Bisarya, Gurunathan and Wheelwright 2003) was used to assess systemizing. The original Systemizing Quotient was a self-report questionnaire with 60 items (including 20 filler items). Items in the original Systemizing Quotient were drawn primarily from traditionally male domains. To improve Systemizing Quotient in this respect, a new revised version (SQ-R) was developed (Wheelwright et al. 2006) that includes new items that are relevant to females in the general population. Systemizing Quotient Revised has high reliability and validity indices (Wheelwright et al. 2006), with α Cronbach equalling .90. For the purpose of the present research a Polish version of the 75-item Systemizing Quotient Revised was prepared. The translation was checked by the same bilingual specialist in psychology. A 4-point Likert-type scale was used in the questionnaire. Participants received 1 or 2 points for the "systemizing" response (slightly agree and strongly agree, respectively) and 0 for the two other responses. The Cronbach α for the Polish version of Systemizing Quotient Revised was .83.

The last instrument used in the study was Paulhus' Balanced Inventory of Desirable Responding (Paulhus 1984). It is a self-report measure of a tendency to provide positive self-descriptions with satisfactory reliability and validity indices (Paulhus 1984, 1989). The questionnaire captures two distinct aspects of socially desirable responding: a more implicit tendency, i.e. self-deceptive enhancement (SDE, 20 items), and a more conscious one, i.e. impression management (IM, 20 items). The reliability of the Polish version of the questionnaire was established in a previous study by Niedźwieńska and Neckar (2013) and was satisfactory: Cronbach's α were .79, .70, and .69 for the entire scale, the SDE subscale, and the IM subscale respectively. The three questionnaires were administered during the same session and their administration was counterbalanced.

RESULTS

DESIGN

The design was a simple factorial design, with gender as a between-subjects factor and desirable responding as a covariate.

RESULTS AND DISCUSSION

For all statistical tests reported below, the rejection level was set at .05 (unless otherwise specified).

To assess gender differences in empathy, a one-way ANOVA with gender as a factor and Empathy Quotient as a dependent variable was performed. It revealed significant gender differences, $F(1, 91) = 15.81$, $\eta^2_p = .15$, with women scoring higher ($M = 44.35$, $SD = 9.56$) compared to men ($M = 36.94$, $SD = 6.46$). Next, two one-way ANCOVAs were conducted to determine a difference between women and men on Empathy Quotient while controlling for Self-Deceptive Enhancement and Impression Management. There was a significant effect of gender on the Empathy Quotient after controlling for Self-Deceptive Enhancement, $F(1, 90) = 16.35$, $\eta^2_p = .15$, and Impression Management, $F(1, 90) = 13.98$, $\eta^2_p = .13$. Women scored higher compared to men, with controlling for Self-Deceptive Enhancement ($M = 44.41$, $SD = 9.56$; $M = 36.84$, $SD = 6.46$, for women and men respectively) and Impression Management ($M = 44.26$, $SD = 9.56$; $M = 37.10$, $SD = 6.46$, for women and men respectively).

In the following step, gender differences in systemizing were assessed. First, a one-way ANOVA with gender as a factor and the Systemizing Quotient Revised as a dependent variable was calculated, and it revealed no gender differences, $F(1, 91) = .86$, $p = .36$, $\eta^2_p = .01$. Next, the effect of desirable responding on the Systemizing Quotient Revised was established by conducting two one-way ANCOVAs with Self-Deceptive Enhancement and Impression Management scores as covariates. In neither analysis was the gender effect significant: $F(1, 90) = .42$, $p = .58$, $\eta^2_p = .005$ for Self-Deceptive Enhancement, and $F(1, 90) = 2.50$, $p = .12$, $\eta^2_p = .03$ for Impression Management.

Summing up, we found the expected gender differences in empathizing (medium effect size), with women outscoring men. However, the influence of Impression Management on gender differences in empathizing was negligible, i.e. the effect size for gender differences was still medium when controlling for Impression Management. Similarly, the influence of Self-Deceptive Enhancement on gender differences in empathizing was none, i.e. there was the same effect size in gender differences with and without controlling for Self-Deceptive Enhancement. We could not analyse any influence of socially desirable responding on gender differences in systemizing, as we found no gender differences on this dimension at all.

STUDY 2

Study 2 investigated the effect of the activation of gender stereotype on gender differences in the questionnaire and the ability test that measured empathizing-systemizing. In one condition the questionnaire and the ability test were administered without any special

instructions in both the female and male groups. In the other condition the stereotype of women being better at emotion recognition and men being better at technical problems was activated just before administering the questionnaire and the ability test. We hypothesized that gender differences would be larger after the activation of gender stereotype. Specifically, we expected that the stereotype activation would increase the results of women in empathizing and the results of men in systemizing. For the questionnaires the possible mechanism would be the increased tendency to depict oneself in accordance with gender stereotype. For the ability tests the change in performance would reflect the increased motivation to do well on a task that fits the stereotype well.

METHOD

DESIGN

The design was a 2 x 2 factorial design, with group (experimental vs. controls) and gender as between-subject factors.

PARTICIPANTS

A total of 104 participants (52 females) were recruited and randomly assigned to the experimental and control conditions. There were 24 women and 26 men in the experimental condition. They were all students of various humanities and engineering faculties of several universities in Southern Poland. Members of the humanities and engineering faculties were equally distributed across gender as well as across control and experimental groups. The mean age of the participants was 23.08, with standard deviation of 2.22.

INSTRUMENTS

Two of the instruments used in Study 2 were the same as in Study 1: Empathy Quotient and Systemizing Quotient Revised. We calculated reliability for each of them in Study 2. Reliability of Empathy Quotient was similar to that of Study 1: Cronbach $\alpha = .82$. Reliability of Systemizing Quotient Revised was higher, Cronbach $\alpha = .90$, compared to Study 1.

To measure empathy more objectively, the Reading the Mind in the Eyes test (henceforth, Eyes test, Baron-Cohen et al. 2001) was also administered. This is a 36-item measure of human capacity to decipher a mental state from a picture of the eyes region. It has been designed as an advanced measure of mind-reading. The Eyes test uses four mental-state descriptions consisting of the correct target word and three incorrect foil words. The translations of all descriptions were checked by a bilingual (English and Polish) specialist in psychology. The reliability of scale was not originally reported by Baron-Cohen. In the present study Cronbach α equalled .58, which is very consistent with results obtained in other studies (Harkness, Jacobson, Duong and Sabbagh 2010; Voracek and Dressler 2006). The analysis of items showed that the accuracy of responses was more than 50% for most items, and only 5 items had a response rate between 38% to 44%. This confirms that target words were chosen at higher rates than expected by chance.

The Folk Physics Test (Baron-Cohen, Wheelwright, Spong, Scahill, and Lawson 2001) was administered to measure skills of systemizing. The test comprises 20 items that depict

various mechanisms reflecting the laws of physics. The items take the form of questions about what will happen next, with the answers in the multiple-choice format. In their original paper the authors of the test provided only information on accurate responses on each item. In the present study we calculated reliability of this test and Cronbach alpha = .52, which is a rather low value. However, the percentage of correct responses was similar to that obtained by Baron-Cohen and his collaborators (2001), with the correct performance of over 50% on the majority of items. Performance was lower (23% to 45%) on seven items only.

PROCEDURE

The questionnaires and tests were administered in small groups, which were gender specific. The first participants were given Empathy Quotient and Systemizing Quotient Revised, which was followed by the administration of the Eyes test and Intuitive Physics test. In the control condition participants were given the following instructions: "Our research concerns how different aspects of personality are related to an ability to solve problems. You will be given two personality questionnaires and two ability tests that measure cognitive skills in solving problems in various domains." In the experimental condition the instructions read: "Our research concerns how different aspects of personality are related to an ability to solve problems. You will be given two personality questionnaires and two ability tests that measure cognitive skills in solving problems in various domains. If you encounter any problems in solving the problems please do not take it personally, because women are known to be much better in recognizing emotions of others, whereas men are much better in solving technical problems."

RESULTS AND DISCUSSION

The Empathy Quotient responses were entered into a 2 (group: stereotype activated vs no activation) x 2 (gender) factorial ANOVA. The main effect of gender was significant $F(1, 100) = 9.37$, $\eta^2_p = .09$, with female participants scoring higher than males. Also, participants in the experimental condition scored higher in empathy than the controls, $F(1, 100) = 6.29$, $\eta^2_p = .06$. The interaction was not significant, $F(1, 100) = 1.12$, $p = .29$, $\eta^2_p = .01$. To test whether, as predicted, women in the experimental condition scored higher compared to women in the control condition we performed a simple effects analysis which revealed a significant difference $F(1, 100) = 6.33$, $\eta^2_p = .06$. There was no significant difference between experimental and control conditions for men. We also calculated a simple effects analysis for gender differences in the control and experimental groups to test the hypothesis that gender differences would be larger in the latter groups. Women scored higher compared to men in the experimental group $F(1, 100) = 8.16$, $\eta^2_p = .08$. For the control condition the difference did not approach statistical significance, $F(1, 100) = 2.09$, $p = .15$, $\eta^2_p = .02$ (see Figure 1).

Similarly, the Systemizing Quotient responses were entered into a 2 (group) x 2 (gender) factorial ANOVA. No significant effects were found, all *p*s > .10. There was however a gender difference in the experimental group that approached statistical significance, $F(1, 100) = 3.70$, $p = .057$, $\eta^2_p = .04$, with women scoring lower than men.

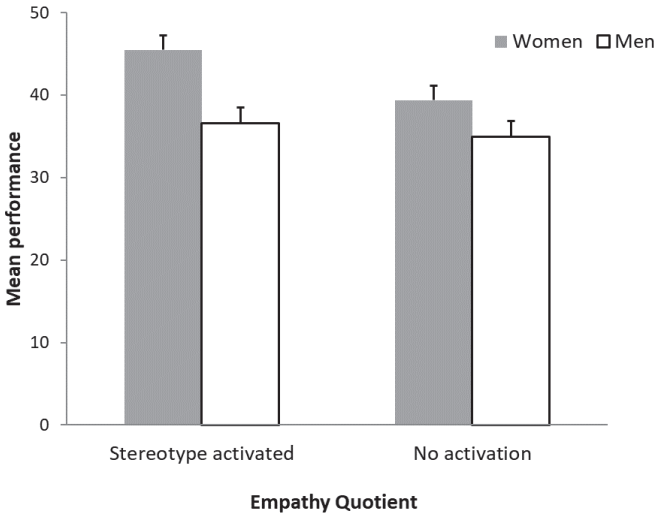

Figure 1. Mean scores (with standard error of the mean) of men and women in Empathy Quotient as a function of stereotype activation

Next, we examined the performance in the Eyes test. The main effect of gender was significant, $F(1, 100) = 13.96$, $\eta^2_p = .12$. As expected, female participants scored higher ($M = 24.31$; SD = 4.13) compared to males ($M = 21.35$; $SD = 3.89$) in the test. The main effect of group and the interaction were not significant ($F_s < 1$).

To examine differences in performance on the Intuitive Physics test, the scores were entered into a 2 (group) x 2 (gender) factorial ANOVA. There was a main effect of gender, $F(1, 100) = 21.44$, $\eta^2_p = .18$, reflecting better performance of male participants. The main effect of group was not significant ($F < 1$). However, this was qualified by a two-way interaction between gender and group, $F(1, 100) = 7.34$, $\eta^2_p = .07$. As expected, male participants scored higher than females in the experimental condition, $F(1, 100) = 25.94$, $\eta^2_p = .21$, but not in the control condition ($p = .17$). Men in the experimental condition scored higher than men in the control condition, $F(1, 100) = 5.24$, $\eta^2_p = .05$. The condition did not have a significant effect for women, $F(1, 100) = 2.38$, $p = .13$, $\eta^2_p = .02$ (see Figure 2).

The expected pattern of gender differences was partially supported in Study 2. As we expected, women in the stereotype activation condition scored higher compared to men, and also higher compared to women in the control condition when empathizing was measured by the questionnaire. The effect of stereotype activation was not observed in the Eyes test. However, the test revealed significant gender differences overall (medium size effect).

Interestingly, a different pattern was revealed when the questionnaire and the ability test were used to measure systemizing. When measured by the questionnaire, systemizing showed only a weak gender difference in the experimental condition. In contrast, our hypotheses were fully confirmed when systemizing was measured by the ability test, i.e. we found stronger gender differences in the experimental condition as well as better performance of men in the experimental group compared to men in the control group.

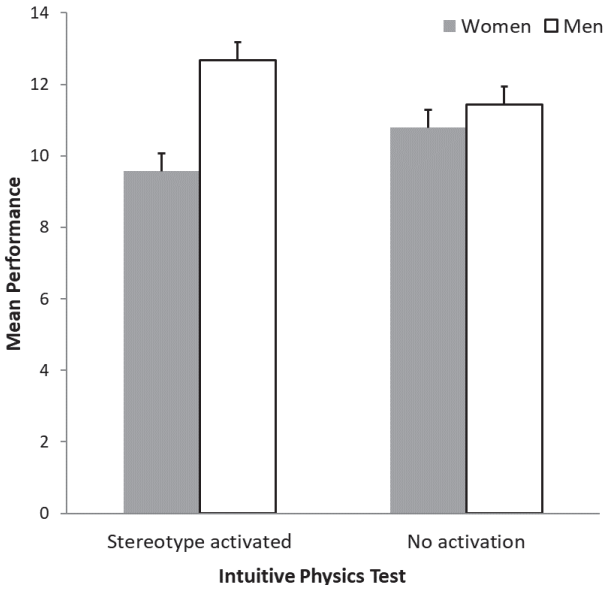

Figure 2. Mean scores (with standard error of the mean) of men and women in Intuitive Physics test as a function of stereotype activation

GENERAL CONCLUSIONS

The present studies examined socio-cultural effects on gender differences in empathizing-systemizing. Generally, our studies add to previous reports that showed superiority of females in empathizing as measured by questionnaires and ability tests. Critically, our results extend previous findings by showing the gender stereotype effects on gender differences in empathizing-systemizing that were particularly strong in the domain of systemizing.

First, our findings generally support the effect size of gender differences in empathizing that has been reported in previous studies (see Groen et al. 2015). Consistently across our two studies and the two measurement methods used (the questionnaire and the ability test), women scored higher and the size of the effect was medium. The pattern of results related to systemizing was inconclusive. Although the ability test revealed large gender differences in favour of men, the systemizing questionnaire showed only a weak gender difference in the gender stereotype activation group. A probable cause of weak gender effect in the questionnaire scores might have been the use of the revised version of the Systemizing Quotient, which favours men to a lesser extent compared to the original version of the systemizing questionnaire (Wheelwright et al. 2006).

Second, socially desirable responding had no effect on the size of gender differences in empathizing. As we did not reveal gender differences in systemizing in Study 1, the possible influence of socially desirable responding on systemizing needs further data collection.

This pattern of results, i.e. an effect size for empathizing similar to other studies on European (Groen et al. 2015; Preti et al. 2012) and Asian (Kim and Lee 2010) samples and the unchanged effect size while controlling for socially desirable responding, seems to be in accordance with the hypothesis put forward by Baron-Cohen (2009) according to which gender differences in empathizing are biologically determined.

However, the results of Study 2 suggest that gender differences in empathizing-systemizing are not immune to socio-cultural influences. The effect of gender stereotype activation was strong in the empathizing questionnaire and in the Intuitive Physics test. The former probably shows a self-presentational effect of gender stereotype activation on empathizing, which should be taken into account in future research.

There was no effect of stereotype activation on performance in the Eyes test, whereas this influence was strong and in the expected direction in the Intuitive Physics test. This somewhat perplexing pattern of results is nevertheless coherent with the results of many studies that investigated the influence of gender stereotypes on the performance in the female and male groups (Grand, Ryan, Schmitt and Hmurovic 2011). Women seem to be more susceptible to stereotype threat, whereas men seem to be more immune to stereotype threat. Therefore, women tend to perform below their actual abilities on the tasks in which men are presumably better, whereas performance of men on the tasks in which women are presumably better is not affected by the awareness of a gender difference in performance (Franceschini, Galli, Chiesi and Primi 2014; Seibt and Förster 2004). The results of Study 2 suggest that men were especially motivated to perform well in the Intuitive Physics task when their gender stereotype was activated. A possible mechanism that may have contributed to their better performance compared to the control group could be linked to interpreting the test situation as a challenge rather than a threat. If a person is more motivated while performing the Intuitive Physics test there is a greater chance of success, as the items of the test require focusing and thinking hard to find a solution to them. The better performance of men on the Intuitive Physics test under the gender stereotype activation is congruent with existing research on the stereotype boost, which is the result of exposure to positive stereotypes (Shih et al. 2002). There are specific mechanisms that may underlie this performance boost, i.e. reducing anxiety, increasing efficiency in neural processing, and activating ideomotor processes (Shih, Pittinsky and Ho 2012).

No effect of stereotype activation on the performance of women in the Eyes test could be explained by a greater difficulty of the Eyes test. Overall, women scored higher than men in this test and it is likely that even if they experienced stereotype boost, they were not able to significantly improve their performance. The Eyes test requires more intuitive knowledge and thus it may be that women could not profit so much from the stereotype boost. The participant just needs to possess the ability to recognize what emotion a person in the picture is expressing. An alternative explanation of lack of stereotype activation effect on women would refer to observed better effects of stereotype boost when stereotype activation was in a more implicit than explicit form (Shih, Pittinsky and Ho 2012).

Some limitations of the present studies should be acknowledged. The samples were rather small, which may explain why some of the effects were not significant. However, the effect size of gender differences in empathizing was similar to those obtained in other studies with bigger samples (Groen et al. 2015; Preti et al. 2011). As the participants were university

students, it limits the interpretation of results to young adults. Also, some of the measures used were ability tests but their reliability was rather low, which might have attenuated some of the effects.

Our findings suggest that gender stereotype activation can be a serious threat to the validity of results obtained by measuring the empathizing-systemizing dimension in the female and male groups. As knowledge of gender differences in empathizing becomes more and more common knowledge, special care should be taken not to activate gender stereotypes in the context of research in which an empathizing questionnaire is used. In addition, similar precautions are recommended in the case of the Intuitive Physics test of systemizing. These caveats apply specifically in those clinical settings where the instruments of this study would be used as a part of autism spectrum diagnosis.

In conclusion, we found robust gender differences in empathizing that were not influenced by socially desirable responding. However, these differences were not entirely immune to social influences: the activation of the gender stereotype made respondents, especially women, present themselves as more empathetic persons. In addition, we found gender differences in systemizing, as measured by the ability test, that were also influenced by the gender stereotype activation. The stereotype activation improved the performance of men, which may be attributed to the stereotype boost, i.e. increasing their motivation to do well in those tests that fitted the male stereotype well.

REFERENCES

Baldner, Conrad and Jared J. McGinley. 2014. *Correlational and exploratory factor analyses (EFA) of commonly used empathy questionnaires: New insights*, "Motivation and Emotion", 38, 5: 727–744.

Baron-Cohen, Simon. 2002. *The extreme male brain theory of autism*, "Trends in Cognitive Sciences", 6: 248–254.

Baron-Cohen, Simon. 2003. *The essential difference: Men, women and the extreme male brain*, London: Penguin.

Baron-Cohen, Simon. 2004. *The cognitive neuroscience of autism*, "Journal of Neurology, Neurosurgery and Psychiatry", 75: 945–948.

Baron-Cohen, Simon. 2006. *The hyper-systemizing, assortative mating theory of autism*. "Progress in Neuropsychopharmacology and Biological Psychiatry", 30: 865–872.

Baron-Cohen, Simon. 2007. *The evolution of empathizing and systemizing: assortative mating of two strong systemizers and the cause of autism*, in: Robin I.M. Dunbar and Louise Barrett (eds.), *The Oxford handbook of evolutionary psychology*, Oxford: Oxford University Press, pp. 213–226.

Baron-Cohen, Simon. 2009. *Autism: the empathizing-systemizing (E-S) theory*, in: Alan Kingstone and Michael Miller (ed.), *The Year in Cognitive Neuroscience*, "Annals of the New York Academy of Science", 1156: 68–80.

Baron-Cohen, Simon. 2010. *Empathizing, systemizing, and the extreme male brain theory of autism*. "Progress in Brain Research", 186: 167–175.

Baron-Cohen, Simon, Rebecca Knickmeyer, and Matthew Belmonte. 2005. *Sex differences in the brain: implications for explaining autism*, "Science", 310: 819–823.

Baron-Cohen, Simon, Jennifer Richler, Dheraj Bisarya, Nhishanth Gunathan and Sally Wheelwright. 2003. *The Systemizing Quotient (SQ): An investigation of adults with Asperger Syndrome or High Functioning Autism, and normal sex differences*, "Philosophical Transactions of the Royal Society of London", 358: 361–374.

Baron-Cohen, Simon, Angel Riviere, Pippa Cross, Massato Fukushima, Catherine Bryant, Maria Sotillo, Julie Hadwin and Dovina French. 1996. *Reading the mind in the face: a cross-cultural and developmental study*, "Visual Cognition", 3: 39–59.

Baron-Cohen, Simon and Sally Wheelwright. 2004. *The Empathy Quotient (EQ). An investigation of adults with Asperger Syndrome or High Functioning Autism, and normal sex differences*, "Journal of Autism and Developmental Disorders", 34: 163–175.

Baron-Cohen, Simon, Sally Wheelwright, Jacqueline Hill, Yogini Raste and Ian Plumb. 2001. *The 'Reading the Mind in the eyes' test revised version: A study with normal adults, and adults with Asperger Syndrome or High Functioning Autism*, "Journal of Child Psychology and Psychiatry", 42: 241–252.

Baron-Cohen, Simon, Sally Wheelwright, Victoria Scahill, Amanda Spong and John Lawson. 2001. *Studies of theory of mind: are intuitive physics and intuitive psychology independent?*, "Journal of Developmental and Learning Disorders", 5: 47–78.

Bar-On, Reuven. 2004. *The Bar-On Emotional Quotient Inventory (EQ-i): Rationale, description and summary of psychometric properties*, in: Glenn Geher (ed.), *Measuring emotional intelligence: Common ground and controversy*, New York: Nova Science, pp. 115–145.

Bateson, Daniel. 2009. *These things called empathy: Eight related but distinct phenomena*, in: Jean Decety and William Ickes (eds.), *The social neuroscience of empathy*, Cambridge, MA: The MIT Press, pp. 3–15.

Berthoz, Sylvie, Michele Wessa and Gayannee Kedia, Bruno Wicker and Julie Grèzes. 2008. *Cross-cultural validation of the empathy quotient in a French-speaking sample*, "The Canadian Journal of Psychiatry / La Revue Canadienne de Psychiatrie", 53, 7: 469–477.

Block, Jack. 1965. *The challenge of response sets*, New York: Appleton-Century-Crofts.

Cabello, Rosario, Miguel A. Sorrel, Irene Fernandez-Pinto, Natlio Extremera, Pablo Fernandez-Berrocal. 2016. *Age and gender differences in ability emotional intelligence in adults: A cross-sectional study*, "Developmental Psychology", 52: 1486–1492.

Chakrabarti, Bhismadev, Edward Bullmore and Simon Baron-Cohen. 2006. *Empathizing with basic emotions: common and discrete neural substrates*, "Social Neuroscience", 1: 364–384.

Crowne, Douglas P. and David Marlowe. 1960. *A new scale of social desirability independent of psychopathology*, "Journal of Consulting Psychology", 24: 349–354.

Davis, Mark H. 1980. *Individual differences in empathy: A multidimensional approach*, "Dissertation Abstracts International", 40: 3480.

De Waal, Frans B.M. 2008. *Putting the altruism back into altruism: The evolution of empathy*, "Annual Review of Psychology", 59: 279–300.

De Waal, Frans B.M. 2009. *The age of empathy: Nature's lessons for a kinder society*, New York: Harmony Books.

Franceschini, Gulia, Silvia Galli, Francesca Chiesi and Catherina Primi. 2014. *Implicit gender–math stereotype and women's susceptibility to stereotype threat and stereotype lift*, "Learning and Individual Differences", 32: 273–277.

Goldenfeld, Nigel, Baron-Cohen, Simon, Wheelwright, Sally. 2005. *Empathizing and systemizing in males, females, and autism*, "Clinical Neuropsychiatry", 2: 338–345.

Grand, James A., Ann Marie Ryan, Neal Schmitt and Jillian Hmurovic. 2011. *How far does stereotype threat reach? The potential detriment of face validity in cognitive ability testing*, "Human Performance", 24: 1–28.

Groen, Yvonne, Anselm M. Fuermaier, Den A.E. Heijer, Olivier Tucha and Monika Althaus. 2015. *The Empathy and Systemizing Quotient: The psychometric properties of the Dutch version and a review of the cross-cultural stability*, "Journal of Autism and Developmental Disorders", 45: 2848–2864.

Harkness, Kate L., Jill A. Jacobson, David Duong and Mark A. Sabbagh. 2010. *Mental state decoding in past major depression: Effect of sad versus happy mood induction*, "Cognition and Emotion", 24: 497–513.

Iacoboni, Marco. 2009. *Mirroring people*, New York, NY: Picador.

Kim, JuHee, Lee, Seung J. 2010. *Reliability and validity of the Korean version of the Empathy Quotient Scale*, "Psychiatry Investigation", 7: 24–30.

Knickmeyer, Rebeca, Simon Baron-Cohen, Peter Raggatt and Kevin Taylor. 2006. *Fetal testosterone and empathy*, "Hormones and Behaviour", 49: 282–292.

Lawrence, Emma J., Philip Shaw, Dawn Baker, Simon Baron-Cohen and Anthony S. David. 2004. *Measuring empathy: reliability and validity of the Empathy Quotient*, "Psychological Medicine", 34: 911–919.

Niedźwieńska, Agnieszka and Jacek Neckar. 2013. *Pozytywna tendencyjność w pamięci źródła informacji na własny temat (Positivity bias in memory of the source of self-information)*, "Psychologia Społeczna", 8: 135–139.

Paulhus, Delroy L. 1984. *Two-component models of socially desirable responding*, "Journal of Personality and Social Psychology", 46: 598–609.

Paulhus, Delroy L. 1986. *Self-deception and impression management in test responses*, in: Alois Angleitner and Jerry S. Wiggins (eds.), *Personality assessment via questionnaire*, New York: Springer, pp. 142–165.

Paulhus, Deyroy L. 1989. *Balanced Inventory of Desirable Responding (BIDR-6)*, Preliminary unpublished scale, University of British Columbia, Vancouver, Canada.

Paulhus, Deyroy L. and Douglas B. Reid. 1991. *Enhancement and denial in socially desirable responding*, "Journal of Personality and Social Psychology", 60, 2: 307–317.

Petrides, Konstantinos V. and Adrian Furnham. 2001. *Trait emotional intelligence: Psychometric investigation with reference to established trait taxonomies*, "European Journal of Personality", 15: 425–448.

Preti, Antonio, Marcello Vellante, Zucca Baron-Cohen, Giulia Simon, Donatella R. Petretto, Carmelo Masala. 2011. *The Empathy Quotient: A cross-cultural comparison of the Italian version*, "Cognitive Neuropsychiatry", 16: 50–70.

Roth, David L., Robert N. Harris, Charles R. Snyder. 1988. *An individual differences measure of attributive and repudiative tactics of favorable self-presentation*, "Journal of Social and Clinical Psychology", 6: 159–170.

Sackeim, Harold A. and Ruben C. Gur. 1978. *Self-deception, self-confrontation, and consciousness*, in: Gary E. Schwartz and David Shapiro (eds.), *Consciousness and self-regulation: Advances in research, Vol. 2*, New York: Plenum Press, pp. 139–197.

Salovey, Peter and John D. Mayer. 1990. *Emotional intelligence*, "Imagination, Cognition and Personality", 9: 185–211.

Schmader, Toni, Michael Johns and Chad Forbes. 2008. *An integrated process model of stereotype threat effects on performance*, "Psychological Review", 115: 336–356.

Seibt, Beate and Jens Förster. 2004. *Stereotype threat and performance: How self-stereotypes influence processing by inducing regulatory foci*, "Journal of Personality and Social Psychology", 87: 38–56.

Shih, Margaret, Nalini Ambady, Jennifer A. Richeson, Kentaro Fujita and Heather M. Gray. 2002. *Stereotype performance boosts: The impact of self-relevance and the manner of stereotype activation*, "Journal of Personality and Social Psychology", 83: 638–647.

Shih, Margaret, Todd L. Pittinsky and Geoffrey C. Ho. 2011. *Stereotype boost: Positive outcomes from the activation of positive stereotypes*, in: Michael Inzlicht and Toni Shamder (eds.), *Stereotype Threat: Theory, Process and Application*, Oxford: Oxford University Press, pp. 141–157.

Simpson, Jeffry and Lorne A. Campbell. 2016. *Methods of evolutionary science*, in: David Buss (ed.), *The handbook of evolutionary psychology*, 2nd ed., vol. 1, Hoboken, NJ: Wiley, pp. 115–135.

Steele, Claude M. 1997. *A threat in the air: How stereotypes shape intellectual identity and performance*, "American Psychologist", 52: 613–629.

Steele, Claude M., Aronson, Joshua. 1995. *Stereotype threat and the intellectual test performance of African Americans*, "Journal of Personality and Social Psychology", 69: 797–811.

Trent, Natalie L., Park, Chamno, Bercovitz, Katherine, Chapman, Ian M. 2016. *Trait socio-cognitive mindfulness is related to affective and cognitive empathy*, "Journal of Adult Development", 23, 62–67.

von Hippel, William., Countrey von Hippel, Leanne Conway, Kristopher J. Preacher, , Jonathan W. Schooler, Gabriel A. Radvansky. 2005. *Coping with stereotype threat: Denial as an impression management strategy*, "Journal of Personality and Social Psychology", 89: 22–35.

Voracek, Martin, Dressler, Stefan G. 2006. *Lack of correlation between digit ratio (2D:4D) and Baron-Cohen's 'Reading the Mind in the Eyes' test, empathy, systemising, and autism-spectrum quotients in a general population sample*, "Personality and Individual Differences", 41: 1481–1491.

Wakabayashi, Akio, Simon Baron-Cohen, Tokyo Uchiyama, Yuko Yoshida, Miho Kuroda and Sally Wheelwright. 2007. *Empathizing and systemizing in adults with and without Autism Spectrum Conditions: cross-cultural stability*, "Journal of Autism and Developmental Disorders", 37: 1823–1832.

Wheeler, Christian S. and Richard E. Petty. 2001. *The effects of stereotype activation on behavior: A review of possible mechanisms*, "Psychological Bulletin", 127: 797–826.

Wheelwright, Sally, Simon Baron-Cohen, Nigel Goldenfeld, Joe Delaney, Debra Fine, Richard Smith, Leonora Weila and Akio Wakabayashi. 2006. *Predicting Autism Spectrum Quotient (AQ) from the Systemizing Quotient-Revised (SQ-R) and Empathy Quotient (EQ)*, "Brain Research", 1079: 47–56.

Wiggins, Jerry S. 1964. *Convergences among stylistic response measures from objective personality tests*, "Educational and Psychological Measurement", 24: 551–562.

Zeyer, Albert, Ayla, Çetin-Dindar, Ahmad N. Md Zain, Mojca Juriševič, Iztok Devetak and Freia Odermatt. 2013. *Systemizing: A cross-cultural constant for motivation to learn science*, "Journal of Research in Science Teaching", 50: 1047–1067.

Zheng, Lijun and Yong Zheng. 2015. *Sex and sexual orientation differences in empathizing-systemizing cognitive styles in China*, "Personality and Individual Differences", 87: 267–271.

RÓŻNICE PŁCIOWE W EMPATYZOWANIU I SYSTEMATYZOWANIU: EFEKT STEREOTYPU PŁCI ORAZ WPŁYW APROBATY SPOŁECZNEJ

Podstawowym celem przedstawionych badań było testowanie hipotez o możliwym wpływie czynników społeczno-kulturowych na istniejące różnice płciowe w zakresie empatyzowania oraz systematyzowania. W pierwszym badaniu kontrolowano efekt zmiennej aprobaty społecznej na różnice płciowe w tych dwóch wymiarach. W badaniu drugim testowano hipotezę o wpływie aktywizowania stereotypu płci na poziom empatyzowania i systematyzowania (w pomiarze zadaniowym oraz samoopisowym). W obu badaniach i obu typach pomiaru kobiety w sposób spójny uzyskały wyniki wyższe od mężczyzn w wymiarze empatyzowania (o umiarkowanej sile efektu). Nie wykazano efektu wpływu zmiennej aprobaty społecznej na wielkość różnic płciowych w zakresie empatyzowania. Jednak aktywizowanie stereotypu płci sprawiło, że kobiety starały się prezentować jako bardziej empatyczne. Dodatkowo aktywizowanie stereotypu płci doprowadziło – wyłącznie u mężczyzn – do polepszenia wykonania zadania mierzącego tendencję do systematyzowania.

Słowa kluczowe: empatyzowanie i systematyzowanie, stereotyp płci, zmienna aprobaty społecznej